Reply to ‘Comment on “Analysis of electroencephalograms in Alzheimer’s disease patients with multiscale entropy”’

J Escudero¹, D Abásolo¹, R Hornero¹, P Espino² and M López¹

¹ Biomedical Engineering Group, E.T.S.I. Telecomunicación, University of Valladolid, Camino del Cementerio s/n, 47011, Valladolid (Spain)

² Hospital Clínico San Carlos, c/Profesor Martín Lagos s/n, 28040, Madrid (Spain)

E-mail: javier.escudero@ieee.org

Abstract

We appreciate the interest of Dr Tang in our article. Certainly, our previous results should be taken with caution due to the small database size. Nevertheless, it must be noted that this limitation was clearly recognised in our article. Furthermore, our hypothesis is completely justified from the current state-of-the-art in the analysis of electroencephalogram (EEG) signals from Alzheimer’s disease (AD) patients. We evaluated whether the multiscale entropy (MSE) analysis of EEG background activity was useful to distinguish AD patients and controls. We do believe that further discussions about risk factors or related clinicophysiological protein aspects are clearly beyond the scope of our article. For the sake of completeness, we now detail some results that complement our previous analysis. They suggest that the MSE analysis can provide relevant information about the dynamics of AD patients’ EEG data. Thus, we must reaffirm our conclusions, although we again acknowledge that further studies are needed.

Keywords: Alzheimer’s disease (AD), electroencephalogram (EEG), multiscale entropy (MSE), receiver operating characteristic (ROC) curve

We appreciate the attention which Dr Tang has paid to our article (Escudero *et al* 2006). Alzheimer’s disease (AD) is a heterogeneous neurodegenerative disorder mainly caused by ageing (Cummings 2004). Additionally, there exist complex interactions of genetic and environmental risk factors (Blennow *et al* 2006). Some of them are associated with vascular disease, including hypercholesterolemia, hypertension, atherosclerosis, coronary heart disease, smoking, obesity and diabetes. However, it is unclear whether these conditions are true causal risk factors or whether they induce cerebrovascular pathology, exceeding the threshold for AD (Blennow *et al* 2006). Although the analysis of these risk factors is relevant to the AD prognosis, we do feel that this discussion clearly exceeds the purpose of our original article. Our objective must not be misunderstood. Rather than being focused on scrutinising either the risk factors and clinicophysiological aspects related to AD or the molecular changes underlying the dementia, our technical article aimed to assess the application of multiscale entropy (MSE) (Costa *et al* 2005) to electroencephalogram (EEG) recordings from AD patients. We wanted to

test whether the MSE might be a useful tool to increase our knowledge of this dementia and to possibly help in its diagnosis (Escudero *et al* 2006).

Actually, the hypothesis tested in our article was that the MSE analysis of EEG background activity might differentiate AD patients from control subjects (Escudero *et al* 2006). This hypothesis is completely justified considering numerous studies that have revealed changes in the non-linear dynamics of EEG recordings from AD patients (for a review, see Jeong 2004). For instance, previous studies by our research group had found a decrease in irregularity and complexity in AD patients' EEG recordings when compared with those of age-matched control subjects (Abásolo *et al* 2005, 2006a, 2006b). Hence, the MSE, which is a non-linear complexity measure recently developed by Costa *et al* (2002), seemed a useful technique to characterise AD patients' EEGs despite the fact that it had not been previously applied to this kind of signals.

The methodology on which the MSE analysis is based is clearly presented in Costa *et al* (2002, 2005), Ferrario *et al* (2006) and Escudero *et al* (2006), among others. Briefly, this analysis is performed by computing the sample entropy on successive coarse-grained versions of the original time series (Costa *et al* 2005, Escudero *et al* 2006). Different features can then be extracted from the MSE profiles. In our study, the slope of the MSE profiles for large time scales provided significant differences between controls and AD patients' EEGs (Escudero *et al* 2006). Furthermore, a recent study has found that this MSE feature also distinguishes control subjects from AD patients when applied to magnetoencephalogram recordings (Gómez *et al* 2007). Although these results are somewhat limited by the reduced number of subjects analysed, they lead us to think that MSE analysis of EEG and MEG data might be helpful to differentiate cognitively normal controls from mild demented patients, as Dr Tang suggests. Nevertheless, further studies should be undertaken to verify this hypothesis.

As we have mentioned, we agree with Dr Tang that our study is limited by the small sample size. This was clearly stated in the "Discussion and Conclusions" section of our article (Escudero *et al* 2006). Due to this limitation, we set a restrictive significance level ($\alpha = 0.01$) to minimize the type I error. We also stated that the study should be extended on a much larger population before MSE derived metrics could be accepted as clinical diagnostic tools (Escudero *et al* 2006).

We applied a Student's *t*-test to evaluate the statistical differences between the MSE slopes of AD patients and control subjects at each electrode. Additionally, receiver operating characteristic (ROC) curves were used to assess the ability of this analysis to discriminate both subject groups (Bradley 1997, Fawcett 2006, Zweig and Campbell 1993). For each electrode with significant differences, a threshold was estimated as the cut-off point at which the highest accuracy was reached. This criterion implies a trade-off between the sensitivity and specificity obtained for that threshold (Bradley 1997). It is always possible to achieve a fixed value of sensitivity by decreasing specificity (and vice versa). Furthermore, both the true negative and the true positive rates should be taken into account when assessing the ability of a parameter to distinguish two subject groups (Zweig and Campbell 1993). Although it is true that the selection of a specific threshold depends on the prevalence of AD patients and on the false-positive and false-negative costs (Bradley 1997, Zweig and Campbell 1993), the optimum accuracy criterion was used in our technical study to measure the ability of the MSE analysis to tell both groups apart. This criterion enabled us to straightforwardly compare the classification achieved with MSE to that obtained in previous studies with other non-linear analysis methods (Abásolo *et al* 2006a, 2006b).

In addition to the particular values of the sensitivity/specificity pair, we also plotted the ROC curves and provided the areas under the ROC curve (AUC) in order to offer a comprehensive picture about the differences between both subject groups (Bradley 1997, Zweig and Campbell 1993). ROC graphs do not depend on class distributions (Fawcett 2006). The AUC can be interpreted as the probability that a randomly selected individual from the control subjects' group has a value smaller than that of a randomly chosen AD patient. This concept is closely related to the Wilcoxon test of ranks, thus providing information about the separation between the values for demented patients and controls (Bradley 1997, Fawcett 2006, Zweig and Campbell 1993). Additionally, a confidence interval (CI) can be computed. We agree with Dr

Tang that a reduced sample size may increase the standard error of the AUC. Nevertheless, it has also been shown that this standard error decreases as the AUC increases (Bradley 1997). Most of the AUC values reported in our article were larger than 0.85 (Escudero *et al* 2006). Furthermore, it can also be assessed whether the AUC is significantly different from 0.5. The rejection of this null hypothesis provides evidence that the test does have the ability to distinguish the groups (Zweig and Campbell 1993). Finally, if the lower bound of the 95% confidence interval for the AUC is higher than 0.5, the information offered by the parameter can be considered useful for classification (Palacios *et al* 2007). For the sake of completeness, we now include the p -value that the AUC is larger than 0.5 and the 95% CI for the ROC curves shown in our article. These results are depicted in Table 1. It can be observed that, despite the small sample size, all AUC values are significantly different from 0.5 (p -value < 0.01), and that the 95% CIs are better than those roughly guessed by Dr Tang in his comment. In five cases (Fp1, T5, T6, P3 and O1), the lower bound of the 95% CI was higher than 0.75.

As Dr Tang has commented, it is difficult to compare multiple analysis techniques. In our article, the MSE analysis was compared with approximate entropy, sample entropy and Lempel–Ziv complexity. It should be noticed that all these non-linear analysis methods were applied to the same database (Abásolo *et al* 2006a, 2006b, Escudero *et al* 2006). This allowed us to straightforwardly compare the statistical differences and the classification results provided by all these methods even though the sample size is small. This comparison suggests that the MSE analysis may characterise the EEG of AD patients better than other non-linear analysis techniques. This might be to the fact that the MSE considers several time scales simultaneously. Some authors have suggested that the simultaneous inspection of multiple time scales could offer relevant information non-available when analysing biomedical signals with other techniques based on the inspection of only one temporal scale, such as approximate entropy (Costa *et al* 2005, Hoyer *et al* 2005, Palacios *et al* 2007).

Additionally, Dr Tang also mentions the problem that may arise when many statistical significance tests are simultaneously performed. In these cases, it may be necessary to correct the computed p -values by a multiple comparison correction strategy, like the Bonferroni correction. This correction prevents the appearance of spurious significant differences between groups (Jobson 1991). It may be necessary when high-density recording equipment is used, like in Gómez *et al* (2007) where comparisons between AD patients and age-matched controls were simultaneously performed at 148 MEG channels. However, the number of channels analysed in our study was limited to 16 (Escudero *et al* 2006). Consequently, we did not apply any correction procedure to the Student's t -test p -values. Additionally, the channels at which significant differences were found were further analysed using ROC curves. It should also be noticed that the application of the well-known, restrictive Bonferroni correction would have modified the p -values obtained in Abásolo (2006a, 2006b) in the same way that those of

Table 1. AUC results for the slope of the MSE profiles for large time scales on the channels at which differences between groups were significant. In addition to the AUC value, the corresponding p -value that the AUC is larger than 0.5 and the 95% CI for the AUC are shown.

Electrode	AUC	p -value	95% Confidence interval	
F3	0.8430	0.006	0.6642	1.0000*
F7	0.8347	0.008	0.6535	1.0000*
Fp1	0.9339	0.001	0.8308	1.0000*
Fp2	0.8512	0.005	0.6821	1.0000*
T5	0.9174	0.001	0.7939	1.0000*
T6	0.9008	0.001	0.7756	1.0000*
P3	0.9174	0.001	0.7998	1.0000*
P4	0.8512	0.005	0.6878	1.0000*
O1	0.9174	0.001	0.7998	1.0000*
O2	0.8760	0.003	0.7315	1.0000*

* The upper bounds of the 95% CI for the AUC that were larger than 1 were rounded down to 1.0000.

(Escudero *et al* 2006) and the results would still be straightforwardly comparable since the same database was analysed. Furthermore, there may be large differences between groups even though the Bonferroni-corrected p -values are not significant (Jobson 1991).

To sum up, we thank Dr Tang for his interest. We agree with Dr Tang that our results should be taken with caution due to the small database size. This limitation was clearly recognised in our article (Escudero *et al* 2006). Nevertheless, our original hypothesis is fully justified taking into account the current state-of-the-art in the analysis of AD patients' EEG recordings (Jeong 2004). We do believe that the goal of our study and its methodology must not be misunderstood, since we do not aim to discuss the risk factors related to AD but to assess the utility of MSE as a technique to analyse EEG recordings. For the sake of completeness, in this document, we have included p -values for the null hypothesis that the AUC are larger than 0.5 and the 95% CI for these parameters. These results, together with those included in Escudero *et al* (2006), suggest that the MSE analysis can provide relevant information about the dynamics of AD patients' EEG recordings. Thus, we do reaffirm our conclusion that AD may be characterized not only by changes in the brain activity on the shortest time scale, but also by an abnormal behaviour on deeper time scales (Escudero *et al* 2006). Nevertheless, we want to recall that the MSE analysis of EEGs cannot yet be applied in AD diagnosis and that further studies with larger databases are needed.

References

- Abásolo D, Hornero R, Espino P, Álvarez D and Poza J 2006a Entropy analysis of the EEG background activity in Alzheimer's disease patients *Physiol. Meas.* **27** 241–53
- Abásolo D, Hornero R, Espino P, Poza J, Sánchez C I and de la Rosa R 2005 Analysis of regularity in the EEG background activity of Alzheimer's disease patients with approximate entropy *Clin. Neurophysiol.* **116** 1826–1834
- Abásolo D, Hornero R, Gómez C, García M and López M 2006b Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and central tendency measure *Med. Eng. Phys.* **28** 315–22
- Blennow K, de Leon M J and Zetterberg H 2006 Alzheimer's disease *Lancet* **368** 387–403
- Bradley A P 1997 The use of the area under the ROC curve in the evaluation of machine learning algorithms *Pattern Recognit.* **30** 1145–59
- Costa M, Goldberger A L and Peng C K 2002 Multiscale entropy analysis of complex physiologic time series *Phys. Rev. Lett.* **89** 068102
- Costa M, Goldberger A L and Peng C K 2005 Multiscale entropy analysis of biological signals *Phys. Rev. E* **71** 021906
- Cummings J L 2004 Alzheimer's disease *New Engl. J. Med.* **351** 56–67
- Escudero J, Abásolo D, Hornero R, Espino P and López M 2006 Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy *Physiol. Meas.* **27** 1091–106
- Fawcett T 2006 An introduction to ROC analysis *Pattern Recognit. Lett.* **27** 861–74
- Ferrario M, Signorini M G, Magines G and Cerutti S 2006 Comparison of entropy-based regularity estimators: application to the fetal heart rate signal for the identification of fetal distress *IEEE Trans. Biomed. Eng.* **53** 119–25
- Gómez C, Hornero R, Abásolo D, Fernández A and Escudero J 2007 Analysis of MEG recordings from Alzheimer's disease patients with sample and multiscale entropies *Proceedings of the of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, ISBN: 1-4244-0788-5, Lyon, France, pp. 6183–6
- Hoyer D, Pompe B, Chon K H, Hardraht H, Wicher C and Zwiener U 2005 Mutual information function assess autonomic information flow of heart rate dynamics at different time scales *IEEE Trans. Biomed. Eng.* **52** 584–92
- Jeong J 2004 EEG dynamics in patients with Alzheimer's disease *Clin. Neurophysiol.* **115** 1490–505
- Jobson J D 1991 *Applied Multivariate Data Analysis. Volume I: Regression and Experimental Design* (New York: Springer-Verlag) pp 409–16
- Palacios M, Friedrich H, Götze C, Vallverdú M, Bayes de Luna A, Caminal P and Hoyer D 2007 Changes of autonomic information flow due to idiopathic dilated cardiomyopathy *Physiol. Meas.* **28** 677–88
- Zweig MH and Campbell G 1993 Receiver-Operating Characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine *Clin. Chem.* **39** 561–77